# Random Probability Measures via Pólya Sequences: Revisiting the Blackwell-MacQueen Urn Scheme

Hemant Ishwaran* and Mahmoud Zarepour†
*Cleveland Clinic Foundation and University of Ottawa*

Sufficient conditions are developed for a class of generalized Pólya urn schemes ensuring exchangeability. The extended class includes the Blackwell-MacQueen Pólya urn and the urn schemes for the two-parameter Poisson-Dirichlet process and finite dimensional Dirichlet priors among others.

## 1. INTRODUCTION

By making use of a remarkably simple generalized Pólya urn scheme, Blackwell and MacQueen ([1]) described an elegant alternate way to approach the Ferguson Dirichlet process ([2]). Let $X_1, X_2, \ldots$ be a sequence of random elements on a complete separable metric space $\mathcal{X}$ defined by

$$\mathbb{P}\{X_1 \in \cdot\} = \mu(\cdot)/\mu(\mathcal{X}) \tag{1}$$

and

$$\mathbb{P}\{X_{i+1} \in \cdot | X_1, \ldots, X_i\} = \mu_i(\cdot)/\mu_i(\mathcal{X}), \qquad i \geq 1, \tag{2}$$

where $\mu_i(\cdot) = \mu(\cdot) + \sum_{j=1}^{i} \delta_{X_i}(\cdot)$ and $\mu$ is a finite non-null measure on $\mathcal{X}$. Blackwell and MacQueen called such a sequence a *Pólya sequence with parameter* $\mu$. They showed that if $X_1, X_2, \ldots$ was such a sequence, then:

(a) $\mu_i(\cdot)/\mu_i(\mathcal{X})$ converges almost surely to a discrete random probability measure $\mu^*$.

(b) $\mu^*$ is the Ferguson Dirichlet process with parameter $\mu$.

(c) Given $\mu^*$, $X_1, X_2, \ldots$ are independent with distribution $\mu^*$.

Result (c) shows that the Blackwell-MacQueen Pólya sequence is exchangeable, while (b) shows that the sequence is an infinite sample from the Dirichlet process. Thus (a) and (b) combined show that the Pólya urn defined by (1) and (2) is a way to draw values from the Dirichlet process. Moreover, (a) shows that the Dirichlet process is the limit for the urn distribution $\mu_i(\cdot)/\mu_i(\mathcal{X})$, thus providing an alternate way to characterize the Dirichlet process. These facts are far more difficult to prove than the contrapositive result which starts with a sample from the Dirichlet process and shows that such a sample must be exchangeable and can be constructed from a Pólya urn. The latter result follows from elementary properties of the Dirichlet process which we now describe. Let $X_1, X_2, \ldots$ be a sequence derived from a Dirichlet process with parameter $\mu$, i.e:

$$\begin{aligned} (X_i|P) &\stackrel{\text{i.i.d.}}{\sim} P, \qquad i = 1, 2, \ldots \\ P &\sim \mu^*. \end{aligned}$$

It was shown in [2] that $\mu^*(\cdot|X_1, \ldots, X_i)$, the posterior for $\mu^*$ based on the first $i$ observations $X_1, \ldots, X_i$, is also a Dirichlet process, but with an updated parameter $\mu_i$ (see [8], Section 3.2, for a proof using a Laplace functional argument). An immediate consequence of this is that

$$\begin{aligned} &\mathbb{P}\{X_{i+1} \in \cdot | X_1, \ldots, X_i\} \\ &= \int \mathbb{P}\{X_{i+1} \in \cdot | X_1, \ldots, X_i, P\}\, \mu^*(dP|X_1, \ldots, X_i) \\ &= \int P(\cdot)\, \mu^*(dP|X_1, \ldots, X_i) \\ &= \mu_i(\cdot)/\mu_i(\mathcal{X}), \end{aligned}$$

and, thus, $X_1, X_2, \ldots$ can be defined by the Pólya urn described by (1) and (2) (the fact that the sequence is exchangeable follows by definition).

Put another way, elementary properties for the Dirichlet process shows us that the *prediction rule*, that is, the conditional distribution for $X_{i+1}$ given $X_1, \ldots, X_i$, corresponds to an exchangeable generalized Pólya urn distribution $\mu_i(\cdot)/\mu_i(\mathcal{X})$. This type of direct result is somewhat unique as it is generally hard to derive simple explicit prediction rules for a general random discrete probability measure. Instead, another way to approach the problem is in the direction studied by Blackwell and MacQueen. Thus, it is natural to wonder what types of Pólya urn schemes other than (1) and (2) lead to: (i) an exchangeable sequence $X_1, X_2, \ldots$ and (ii) an urn distribution with a limiting random discrete probability measure? Sufficient and necessary conditions for (i) have been given in [3] (cf Theorem 2) in terms of what is called the exchangeable partition probability function (EPPF), a symmetric non-negative function which characterizes the distribution of an exchangeable partition on the positive integers $\{1, 2, \ldots\}$. In this paper, however, we take a more direct approach to answering (i) (and consequently (ii)), by introducing an Exchangeability Condition (Section 3) which puts constraints on the manner in which the urn scheme selects new values or chooses a previously sampled value. While these conditions will be shown only to be sufficient to ensure (i), they have the advantage that they are simpler to understand then conditions stated in terms of the EPPF. The proof should also be readily accessible to non-experts to this area. Our Exchangeability Condition is shown to hold for several important generalized Pólya urns, including those for the two-parameter Poisson-Dirichlet process ([11]) as well as the class of finite dimensional Dirichlet priors ([5]). Corollary 1 of Section 3, our main result, summarizes our results.

*Electronic address: ishwaran@bio.ri.ccf.org
†Electronic address: zarepour@mathstat.uottawa.ca

## 2. NOTATION AND BACKGROUND

We begin by introducing some notation necessary to explain our generalized Pólya urn schemes. Let $\mathbf{p}_i = \{C_{j,i} : j = 1, \dots, n(\mathbf{p}_i)\}$ denote a partition of $\{1, \dots, i\}$ where $C_{j,i}$ is the $j$th set of the partition. Write $e_{j,i}$ for the cardinality of $C_{j,i}$. Thus, $\mathbf{p}_i$ is a partition made of $n(\mathbf{p}_i)$ sets and $\sum_{j=1}^{n(\mathbf{p}_i)} e_{j,i} = i$. Let $X_1^*, X_2^*, \cdots$ denote the sequence of unique values *in the order of their appearance* from $X_1, X_2, \dots$ and let $\mathbf{p}_i$ be a partition of $\{1, \cdots, i\}$ recording the clustering of the first $i$ observations $X_1, \dots, X_i$. By this we mean $X_l = X_j^*$ for each $l \in C_{j,i}$, where $j = 1, \dots, n(\mathbf{p}_i)$.

Let $\nu$ denote a non-null probability measure over $\mathcal{X}$. We will consider sequences $X_1, X_2, \dots$ defined by

$$\mathbb{P}\{X_1 \in \cdot\} = \nu(\cdot) \tag{3}$$

and

$$\begin{aligned} &\mathbb{P}\{X_{i+1} \in \cdot | X_1, \dots, X_i\} \qquad (4)\\ &= \frac{q_{0,i}}{\sum_{j=0}^{n(\mathbf{p}_i)} q_{j,i}} \nu(\cdot) + \sum_{j=1}^{n(\mathbf{p}_i)} \frac{q_{j,i}}{\sum_{j=0}^{n(\mathbf{p}_i)} q_{j,i}} \delta_{X_j^*}(\cdot), \qquad i \geq 1, \end{aligned}$$

where

$$q_{j,i} := q_{j,i}(e_{1,i}, \dots, e_{n(\mathbf{p}_i),i}),\ q_{0,i} := q_{0,i}(e_{1,i}, \dots, e_{n(\mathbf{p}_i),i}) \tag{5}$$

are non-negative real valued symmetric functions depending only upon $\{e_{1,i}, \dots, e_{n(\mathbf{p}_i),i}\}$.

The form for $q_{0,j}$ and $q_{j,i}$ in (5) is suggested by Theorem 1 of [3] which states that for $X_1, X_2, \cdots$ to be exchangeable, the functions $q_{0,i}$ and $q_{j,i}$ must be almost surely equal to some function of the partition $\mathbf{p}_i$; or equivalently, they must be some function of the cardinalities $e_{j,i}$. For example, observe that the Blackwell-MacQueen Pólya sequence (with parameter $\mu$) corresponds to the choices $q_{0,i} = \mu(\mathcal{X})$, $q_{j,i} = e_{j,i}$ and $\nu(\cdot) = \mu(\cdot)/\mu(\mathcal{X})$.

In proving our general result, an important technical condition that we will need to address concerns the choice for $\nu$. We say that $\nu$ is *non-atomic* if $\nu\{x\} = 0$ for each $x \in \mathcal{X}$. One of the unique features of the Blackwell-MacQueen Pólya urn scheme is that it yields an exchangeable sequence regardless of whether $\nu$ is non-atomic. For example, if $\mathcal{X} = \{1, \dots, r\}$ is a finite sample space and $\mu$ is a finite discrete measure over $\mathcal{X}$, then (1) and (2) implies that $X_1, \dots, X_i$ is the result of successive draws from an urn originally having $\mu(l)$ balls of color $l$, and following each draw for a ball, the ball is replaced and another one of its same color is added to the urn. It follows that

$$\mathbb{P}\{X_1 = x_1, \dots, X_i = x_i\} = \frac{1}{\mu(\mathcal{X})^{[i]}} \prod_{l=1}^{r} \mu(l)^{[n(l)]},$$

where $n(l)$ denotes the number of $x$'s equal to $l$ and $a^{[i]} = a(a+1)\cdots(a+i-1)$ (note: $a^{[0]} = 1$). Observe that since the right-hand side is a symmetric function of $(x_1, \dots, x_i)$, it follows automatically that $X_1, \dots, X_i$ is exchangeable. This fact was not lost on Blackwell and MacQueen. Indeed, the key to their proof relies on the fact that their Pólya urn scheme produces an exchangeable sequence for finite sample spaces.

### 2.1. Non-exchangeability over discrete spaces

However, prediction rules for random discrete measures often break down when $\mathcal{X}$ is allowed to be a finite sample space. A good example is the two-parameter Poisson-Dirichlet process discussed in [11]. This is the random discrete probability measure whose prediction rule for a non-atomic $\nu$ corresponds to the choices

$$q_{0,i} = \theta + \alpha n(\mathbf{p}_i) \quad \text{and} \quad q_{j,i} = e_{j,i} - \alpha,$$

where $0 \leq \alpha < 1$ and $\theta > -\alpha$. See [11] and also [9], [10] for further details. Setting $\alpha = 0$ and $\theta = \mu(\mathcal{X})$ leads to the Blackwell-MacQueen Pólya sequence with parameter $\mu = \theta\nu$, and as discussed produces an exchangeable sequence without constraint to $\nu$. In general, however, if $\alpha \neq 0$, exchangeability breaks down if $\mathcal{X}$ is allowed to be a finite sample space and $\nu$ is atomic. This can be easily demonstrated by the following counter-example. Let $\mathcal{X} = \{1, \dots, r\}$ where $r \geq 2$ and suppose that $\nu(l) = 1/r$ for each $l = 1, \dots, r$. Then,

$$\begin{aligned} &\mathbb{P}\{X_1 = 1, X_2 = 2, X_3 = 1\} \\ &= \mathbb{P}\{X_1 = 1\} \times \mathbb{P}\{X_2 = 2 | X_1\} \\ &\qquad \times \mathbb{P}\{X_3 = 1, X_3 = X_1, X_3 \neq X_1 | X_1, X_2\} \\ &= \frac{(\theta+\alpha)\Big((\theta+2\alpha)/r + 1 - \alpha\Big)}{r^2(\theta+1)(\theta+2)}. \end{aligned}$$

Note that the last expression in the middle equation underlies the problem with working with an atomic measure. Here the conditional event that $X_3 = 1$ occurs if we choose the previous value $X_1$ or if we choose the value $X = 1$ randomly from $\nu$. This wouldn't be a problem with a non-atomic probability measure since the probability of obtaining a previously observed $X_i$ value is zero under $\nu$. But this leads to a breakdown of exchangeability for an atomic measure. Consider the probability,

$$\begin{aligned} &\mathbb{P}\{X_1 = 1, X_2 = 1, X_3 = 2\} \\ &= \mathbb{P}\{X_1 = 1\} \times \mathbb{P}\{X_2 = 1, X_2 = X_1, X_2 \neq X_1 | X_1\} \\ &\qquad \times \mathbb{P}\{X_3 = 2 | X_1, X_2\} \\ &= \frac{\Big((\theta+\alpha)/r + 1 - \alpha\Big)(\theta+\alpha)}{r^2(\theta+1)(\theta+2)}. \end{aligned}$$

Thus, $\mathbb{P}\{X_1 = 1, X_2 = 2, X_3 = 1\} \neq \mathbb{P}\{X_1 = 1, X_2 = 1, X_3 = 2\}$ unless $\alpha = 0$. This shows that only the Blackwell-MacQueen Pólya urn is exchangeable in general for the two-parameter process.

## 3. MAIN RESULTS

Thus, given the technical difficulties in working with atomic measures, we will hereafter restrict attention to non-atomic measures $\nu$. Our results will also rely on the following key conditions for the functions $q_{0,i}$ and $q_{j,i}$ appearing in (3) and (4).

**Exchangeability Condition.** *For each $i \geq 1$, $q_{j,i} = \psi(e_{j,i})$ and $q_{0,i} = \psi_0(n(\mathbf{p}_i))$, where $\psi$ and $\psi_0$ are some fixed non-negative real valued functions. Furthermore, for each partition $\mathbf{p}_i$ of $\{1, \ldots, i\}$*

$$\sum_{j=0}^{n(\mathbf{p}_i)} q_{j,i} = \xi(i) > 0 \tag{6}$$

*where $\xi$ is some fixed positive real valued function.*

These conditions are satisfied by many interesting generalized Pólya urn schemes, which we now list. By satisfying the Exchangeability Condition, Theorem 1 (stated later) shows that each of these urns (subject to a non-atomic $\nu$) are exchangeable.

1. Independent and identically distributed sampling. This is of course the most obvious form of exchangeability and follows with choices $q_{0,i} = 1$ and $q_{j,i} = 0$.

2. $N$ values selected at random. Let $N > 1$ be a positive integer and let $q_{0,i} = (N - n(\mathbf{p}_i))I\{n(\mathbf{p}_i) < N\}$ and $q_{j,i} = 1$. Observe that $q_{0,i}$ becomes zero when $n(\mathbf{p}_i) \geq N$, which restricts the urn sequence from having more than $N$ distinct values. Note that condition (6) is satisfied because $\sum_{j=0}^{n(\mathbf{p}_i)} q_{j,i} = N$.

3. The Blackwell-MacQueen Pólya sequence with parameter $\mu$. This corresponds to $q_{j,i} = e_{j,i}$, $q_{0,i} = \mu(\mathcal{X})$ and $\sum_{j=0}^{n(\mathbf{p}_i)} q_{j,i} = \mu(\mathcal{X}) + i$.

4. The two-parameter Poisson-Dirichlet process. As discussed, this corresponds to choices $q_{0,i} = \theta + \alpha n(\mathbf{p}_i)$ and $q_{j,i} = e_{j,i} - \alpha$, where $0 \leq \alpha < 1$ and $\theta > -\alpha$. Thus, $\sum_{j=0}^{n(\mathbf{p}_i)} q_{j,i} = \theta + i$. This is the prediction rule for the random discrete probability measure $\mathcal{P}$ defined by

$$\begin{aligned}\mathcal{P}(\cdot) &= V_1 \delta_{Z_1}(\cdot) \\ &+ \sum_{k=2}^{\infty} \{(1 - V_1)(1 - V_2) \cdots (1 - V_{k-1})\, V_k\}\, \delta_{Z_k}(\cdot),\end{aligned} \tag{7}$$

where $\{V_k\}$ are i.i.d Beta$(1 - \alpha, \theta + k\alpha)$ random variables, independent of $\{Z_k\}$, which are i.i.d with law $\nu$. See [11] for details. Observe that by setting $\alpha = 0$ we end up with the Dirichlet process with parameter $\mu = \theta\nu$. In this case, (7) corresponds to the stick-breaking representation for the Dirichlet process. See [4] for background on stick-breaking priors.

5. Finite dimensional Dirichlet priors (Fisher's model). Let $N > 1$ be a positive integer and let $q_{0,i} = \theta(1 - n(\mathbf{p}_i)/N)I\{n(\mathbf{p}_i) < N\}$ and $q_{j,i} = e_{j,i} + \theta/N$, where $\theta > 0$. Then $\sum_{j=0}^{n(\mathbf{p}_i)} q_{j,i} = \theta + i$ which satisfies (6). Observe that the choice for $q_{0,i}$ restricts the process from having more than $N$ distinct values. One can show that the values $q_{0,i}$ and $q_{j,i}$ correspond to the prediction rule for the finite dimensional Dirichlet prior $\mathcal{P}_N$ defined by

$$\mathcal{P}_N(\cdot) = \sum_{k=1}^{N} \frac{G_k}{\sum_{k=1}^{N} G_k}\, \delta_{Z_k}(\cdot),$$

where $\{G_k\}$ are i.i.d Gamma$(\theta/N)$ random variables, independent of $\{Z_k\}$, which are i.i.d with law $\nu$. See [9], [10] and [4] for further details. Also see [6] who showed that $\mathcal{P}_N$ is a weak limit approximation to the Dirichlet process.

### 3.1. Exchangeability

We now show that our Exchangeability Condition is sufficient to ensure that the sequence defined by (3) and (4) is exchangeable.

**Theorem 1.** *If $\nu$ is a non-atomic (and non-null) probability measure over $\mathcal{X}$ and the Exchangeability Condition holds, then $X_1, X_2, \ldots$ is exchangeable.*

*Proof.* Let $i > 1$ (the case $i = 1$ is obvious) and let $dx_1, \ldots, dx_i$ denote a sequence of differentials, some of which can be equal. Let $\mathbf{p}_i = \{C_{j,i} : j = 1, \ldots, n(\mathbf{p}_i)\}$ be the partition of $\{1, \ldots, i\}$ which records the clustering of $dx_1, \ldots, dx_i$. That is, if $dx_1^*, \ldots, dx_{n(\mathbf{p}_i)}^*$ denote the unique values of $dx_1, \ldots, dx_i$, then $dx_l = dx_j^*$ for each $l \in C_{j,i}$. As before, write $e_{j,i}$ for the cardinality of $C_{j,i}$. For notational convenience set $\psi(0) = 1$.

It follows from the assumption that $\nu$ is non-atomic, and upon using (3) and (4), that

$$\begin{aligned}&\mathbb{P}\{X_1 \in dx_1, \ldots, X_i \in dx_i\} \\ &= \prod_{j=1}^{i} \mathbb{P}\{X_j \in dx_j | X_1, \ldots, X_{j-1}\} \\ &= D_i^{-1} \prod_{k=1}^{n(\mathbf{p}_i)-1} \psi_0(k) \prod_{j=1}^{n(\mathbf{p}_i)} \Big\{ \nu(dx_j^*) I\{dx_l = dx_j^* : l \in C_{j,i}\} \\ &\qquad \times \Big( \psi(1) \times \cdots \times \psi(e_{j,i} - 1) \Big) \Big\},\end{aligned} \tag{8}$$

where the first product (in square brackets) follows from the assumption that $q_{0,i} = \psi_0(n(\mathbf{p}_i))$ (note: if $n(\mathbf{p}_i) = 1$ the product is assumed to be 1), while the second product uses the assumption that $q_{j,i} = \psi(e_{j,i})$. The expression $D_i$ appearing in (8) is a normalizing constant. By (6), it can be seen that $D_i = \xi(1) \times \cdots \times \xi(i-1)$. Thus, deduce that the right-hand side of (8) is a symmetric function of $(dx_1, \ldots, dx_i)$, and hence that $X_1, \ldots, X_i$ is exchangeable. $\square$

**Remark 1.** *As a special case, the expression (8) yields the well known joint density for the Dirichlet process:*

$$\begin{aligned}\mathbb{P}\{X_1 \in dx_1, \ldots, X_i \in dx_i\} &= \frac{\mu(\mathcal{X})^{n(\mathbf{p}_i)}}{\mu(\mathcal{X})^{[i]}} \\ \times \prod_{j=1}^{n(\mathbf{p}_i)} \Big( \nu(dx_j^*) I\{dx_l = dx_j^* &: l \in C_{j,i} \Big) (e_{j,i} - 1)!.\end{aligned}$$

*(Substitute $\xi(j) = \mu(\mathcal{X}) + j$, $\psi(j) = j$ and $\psi_0(j) = \mu(\mathcal{X})$ into (8)).*

### 3.2. The Blackwell-MacQueen generalization

By appealing to Proposition 11 of [10], in combination with Theorem 1, we obtain the following corollary which is a generalization of the Blackwell and MacQueen result.

**Corollary 2.** *Let $X_1, X_2, \ldots$ be the sequence defined by (3) and (4) where $\nu$ is a non-atomic (non-null) probability measure over $\mathcal{X}$ and $\{q_{0,i}, q_{j,i}\}$ satisfy the Exchangeability Condition.*

*(a) Let $F_{i+1}$ denote the conditional distribution for $X_{i+1}$ defined by (4). Then, $F_{i+1} \stackrel{\text{a.s}}{\to} \mathcal{P}^*$ in $\ell_1$-distance, where $\mathcal{P}^*$ is the random probability measure defined by*

$$\mathcal{P}^*(\cdot) = \sum_j p_j\, \delta_{X_j^*}(\cdot) + (1 - \sum_j p_j)\nu(\cdot),$$

*where $p_j = \lim_{i\to\infty} e_{j,i}/i$.*

*(b) $\{X_j^*\}$ are i.i.d $\nu$ and independent of $\{p_j\}$.*

*(c) Given $\mathcal{P}^*$, $X_1, X_2, \ldots$ are independent with distribution $\mathcal{P}^*$.*

*(d) If $q_{0,i}/\xi(i) \stackrel{\text{a.s}}{\to} 0$, then $\mathcal{P}^*$ is discrete with probability one; i.e. $\mathcal{P}^*(\cdot) = \sum_j p_j\, \delta_{X_j^*}(\cdot)$.*

*Proof.* Theorem 1 ensures that $X_1, X_2, \ldots$ is exchangeable. Thus, (a), (b) and (c) follows from de Finetti's representation for exchangeable sequences. See Theorem 6 of [9] and Proposition 11 of [10]. To prove (d) we use a theorem of [7] which states that if $X_1, X_2, \ldots$ is an exchangeable sequence from a random measure $\mathcal{P}^*$, then $\mathcal{P}^*$ is discrete with probability one if

$$a_i = \mathbb{P}\{X_{i+1} \text{ is different than } X_1, \ldots, X_i\} \stackrel{\text{a.s}}{\to} 0.$$

See [12], Section 1.6, for a proof. Thus, (d) is proven since $a_i = q_{0,i}/\xi(i)$. □

**Remark 2.** *A little bit of work shows that each of our examples listed earlier (excluding our first example for the i.i.d case) are examples of generalized Pólya urn schemes which satisfy condition (d). Thus, each produce exchangeable sequences from a random discrete probability measure.*


[1] Blackwell, D. and MacQueen, J. B. (1973). Ferguson distributions via Polya urn schemes. *Ann. Statist.*, **1**, 353-355.

[2] Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems *Ann. Statist.*, **1**, 209-230.

[3] Hansen, B. and Pitman, J. (2000). Prediction rules for exchangeable sequences related to species sampling. *Statist. Prob. Letters*, **46**, 251-256.

[4] Ishwaran, H. and James, L. F. (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Statist. Assoc.*, **96**, 161-173.

[5] Ishwaran, H. and Zarepour, M. (2002a). Dirichlet prior sieves in finite normal mixtures. *Statistica Sinica*, **12**, 941-963.

[6] Ishwaran, H. and Zarepour, M. (2002a). Exact and approximate sum-representations for the Dirichlet process. *Canad. J. Statist.*, **30**, 269-283.

[7] Krasker, W. and Pratt, J. W. (1986). Discussion of "On the consistency of Bayes Estimates" by Diaconis and Freedman. *Ann. Statist.*, **14**, 55-58.

[8] Lo, A. Y. and Weng, C. S. (1989). On a class of Bayesian nonparametric estimates: II. Hazard rates estimates. *Ann. Inst. Stat. Math.*, **41**, 227-245.

[9] Pitman, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields*, **102**, 145-158.

[10] Pitman, J. (1996). Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory* (T.S. Ferguson, L.S. Shapley and J.B. MacQueen, eds.) 245-267. IMS Lecture Notes-Monograph series, Vol 30.

[11] Pitman, J. and Yor, M. (1997). The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Prob.*, **25**, 855-900.

[12] Schervish, M. J. (1995). *Theory of Statistics*. Springer, New York.